Distributed Systems

Workshop 1
Semester 1, 2017

Introduction about myself

Xinyi Xu

xinyi.xu@unimelb.edu.au

Linkedin: https://www.linkedin.com/in/xinyi-yvonne-xu-81156a5b/

Graduate project(ongoing): build entity linking based on knowledge graph

Three fun facts:

- Latin dancer 🕺

- Once injured knee during dancing and kept resting for 8 months 😔

- Complete a half-marathon last Oct 🏃

Expectation

- Get prepared for the tutorial

- If not, don't panic

- Come and join the discussion

Structure

- Review lecture content via questions
- Group discussion
- External reading fine points

Q1. Definition of distributed systems

- A system in which hardware or software components located at networked computers communicate and coordinate their actions only by passing message [**Coulouris**]
- A collection of independent computers that appears to its users as a single coherent system [**Tanenbaum**]

Q2. Difference between computer networks and distributed systems

Computer networks - Is a collection of spatially separated, interconnected computers that exchange messages based on specific protocols. Computers are addressed by IP addresses

Distributed systems - Multiple computers on the network working together as a system. The spatial separation of computers and communication aspects are hidden from users.

Q3. Four reasons to use distributed systems

- Economy
    - Sharing hardware, software, processing power etc
- Reliability
    - Fault tolerance, e.g. HTTP error 503
- Availability
    - e.g. 3 machines with 0.95 probability of being up
- Scalability
    - Performance bottleneck of centralized components, e.g. a single mail server

Q4. Issues when use distributed systems

- Concurrency
- No global clock
- Independent failure

External reading

Two examples of why organization would need to build a distributed system:

- The demands of a consumer website/API or multitenant enterprise application simply exceed the computing capacity of any one machine.

- An enterprise moves an existing application, such as a three-tier system, onto a cloud service provider in order to save on hardware/data-center costs
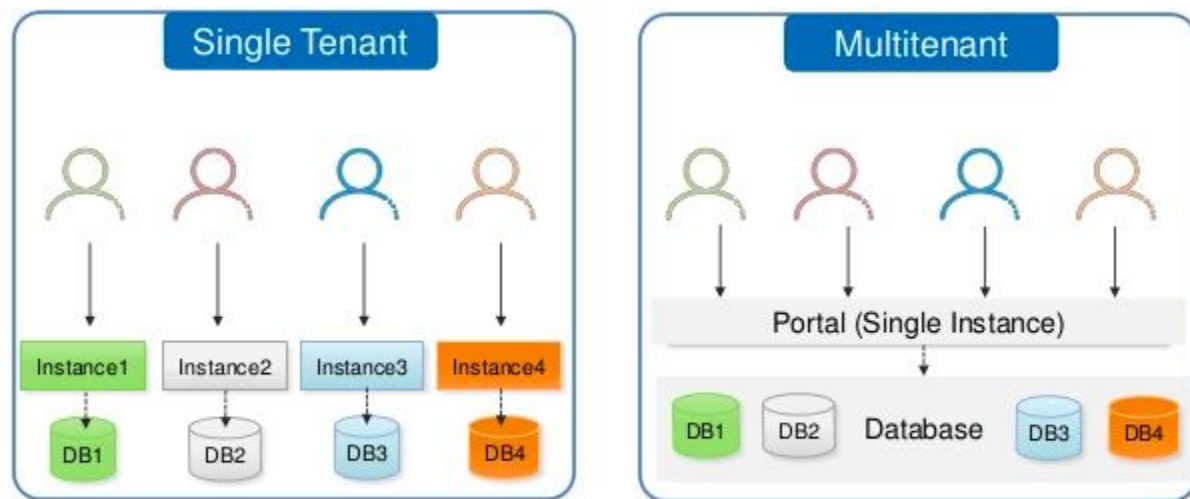
Architecting a Distributed System

Geographies - global or run in 'silos' per region

Data segregation - Single or multi-tenancy model

SLAs(Service Level Agreements) - Availability, latency, throughput, consistency, durability guarantees

Security - IAAA(identity, authentication, authorization, and audit), confidentiality, privacy

Single vs. Multitenant

Image-resizing service

The distributed services of an image resize service.